DOES IMITATION LEARNING PRESERVE EXPERT PERFORMANCE ACROSS TASK EXECUTION SPEEDS? A BENCHMARK STUDY

Anonymous authors
Paper under double-blind review

ABSTRACT

Imitation learning seeks to reproduce expert behavior and is commonly evaluated by task success rate, the fraction of evaluation episodes completed successfully. Matching expert success under nominal conditions does not establish how the expert–learner performance difference changes with task conditions. We study how the expert–learner performance difference varies with execution timing. Timing changes are expressed through a scalar speedup factor; a factor of two halves the scheduled durations. We introduce ParcelStow, a benchmark with three contact-rich manipulation tasks—parcel insertion, upright placement, and keyed peg insertion—that varies these durations while fixing geometry, physical parameters, and success criteria. We evaluate Action Chunking with Transformers (ACT), Diffusion Policy, and DAgger using success rates, task-stage completion, and failure diagnostics. On parcel insertion, ACT matches the expert's 100% at nominal timing but falls to 53% at the maximum demonstrated speedup factor, compared with 84% for the expert. On upright placement and keyed peg insertion, ACT has higher nominal success than the expert, but this ordering reverses beyond the demonstrated speedup factors. These results show that nominal task success does not establish preservation of expert performance under changed task conditions, with execution timing providing a concrete counterexample.

1 INTRODUCTION

In imitation learning, a learner is trained to reproduce expert behavior and is commonly evaluated by task success rate, the fraction of evaluation episodes completed successfully. Matching an expert's nominal success rate does not establish how the expert–learner success-rate difference varies with task conditions. Consider a contact-rich manipulation task in which a robot must lift, reorient, and insert an object. Shortening the allotted execution time increases the scheduled motion speed and can change joint velocities, accelerations, tracking demands, and contact transients. Expert and learner success can therefore respond differently to the same timing requirement.

Existing benchmarks emphasize variation in tasks, scenes, objects, or instructions Yu et al. (2019); James et al. (2019); Mees et al. (2021); Liu et al. (2023); Nasiriany et al. (2024), while recent methods accelerate demonstrations, retime predicted actions, or adapt learned policies for faster execution (Guo et al., 2025; Ranawaka Arachchige et al., 2025; Nam et al., 2026). We evaluate how task-condition perturbations change the expert–learner success-rate difference, using execution timing as the experimental variable. We compare learners whose nominal success matches or exceeds expert success under common timing demands, distinguishing factors represented in the demonstrations from extrapolation beyond them.

ParcelStow implements this comparison on parcel insertion (PARCEL), upright placement (UPRIGHT), and keyed peg insertion (PEG) with a fixed-base humanoid and anthropomorphic hand. We hold geometry, physical parameters, action semantics, and success criteria fixed while varying phase durations (Sec. 3.1). We evaluate ACT (Zhao et al., 2023), Diffusion Policy (DP) (Chi et al., 2023), and DAgger (Ross et al., 2011). The expert–ACT comparisons in Fig. 1 show a 31-percentage-point difference after exact nominal equality on PARCEL, within the demonstrated range, and reversals of nominal ACT advantages on UPRIGHT and PEG beyond their demonstrated ranges. Stage and

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

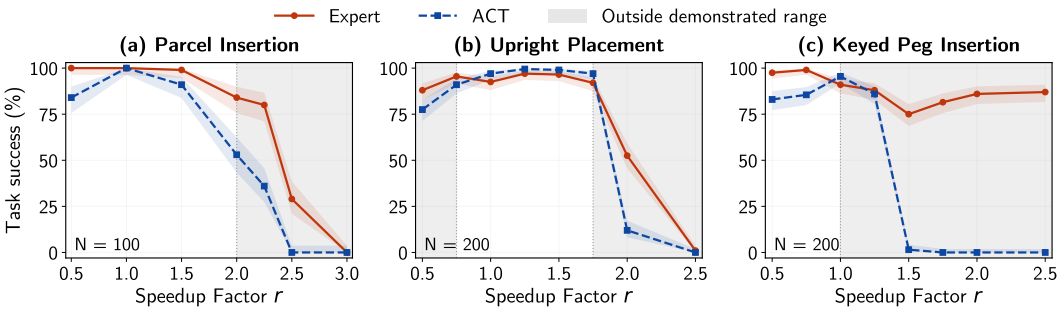


Figure 1: **Expert and ACT Success Across Three Tasks.** The speedup factor r is the ratio of nominal to allotted manipulation-phase duration. Curves show the expert and ACT (Sec. 4.1.4), with pointwise Wilson 95% intervals and 100 episodes per PARCEL factor, or 200 per UPRIGHT and PEG factor. Gray regions mark extrapolation beyond demonstrated speedup factors. On PARCEL, expert and ACT success diverge within the demonstrated range after nominal equality; on UPRIGHT and PEG, their nominal ordering reverses under extrapolation.

handoff analyses localize PARCEL failures to insertion and final orientation (Sec. 5.4). DP provides comparisons from lower nominal success levels; DAgger completes too few nominal episodes for a speed-degradation analysis.

1.1 RELATED WORK

1.1.1 IMITATION AND GENERALIST MANIPULATION

Offline imitation performance varies with demonstration coverage, model design, action representation, and the difference between training and evaluation state distributions (Mandlekar et al., 2021). DAgger addresses the sequential train–test mismatch by aggregating expert labels on states induced by the learner (Ross et al., 2011). BeT and Diffusion Policy model multimodal action distributions (Shafiullah et al., 2022; Chi et al., 2023), PerAct extends behavior cloning to multi-task manipulation through voxelized representations (Shridhar et al., 2022), and ACT predicts multi-step action chunks with temporal ensembling (Zhao et al., 2023). Comparative studies have also examined imitation-learning design choices (Jia et al., 2025), in-context imitation (Vosylius & Johns, 2025), cross-embodiment scaling (Open X-Embodiment Collaboration, 2023; Octo Model Team et al., 2024; Black et al., 2024), and discrepancies between demonstrated and executed motion (Kedia et al., 2024).

Several recent methods explicitly accelerate policy execution. DemoSpeedup trains policies on demonstrations downsampled at segment-specific rates (Guo et al., 2025), while SAIL combines temporally consistent inference, controller-invariant targets, adaptive speed modulation, and latency-aware action scheduling (Ranawaka Arachchige et al., 2025). SpeedAug uses reinforcement learning to adapt a policy initialized from speed-augmented demonstrations (Nam et al., 2026). Reference-phase adaptation combines tracking-error feedback with a learned phase adjustment (Sleiman et al., 2025). Other chunk-based methods accommodate inference latency by asynchronously generating future actions (Black et al., 2025) or adapt execution horizons to determine when new observations enter the policy (Nie et al., 2026). These mechanisms change motion timing, action scheduling, or observation use. We compare a fixed learner and its demonstration expert under a common phase-duration law at speedup factors within and outside the demonstrated range.

1.1.2 DEXTEROUS AND DYNAMIC MANIPULATION

Dexterous manipulation studies have demonstrated learned control of object motion with changing hand–object contacts (Rajeswaran et al., 2017; OpenAI et al., 2018; 2019; Petrenko et al., 2023). Recent systems train grasps for dynamic motion (Zhang et al., 2025), convert static demonstrations into dynamic manipulation (Liao et al., 2026), or select grasps using the downstream motion they must support (Gupta et al., 2026; Enwerem et al., 2026b). These methods seek grasps stable under downstream dynamic motion. Research on grasp quality studies force and form closure (Murray et al., 2017), as well as resistance to external wrenches and task performance through analysis

or simulation (Ferrari & Canny, 1992; Roa & Suárez, 2015; Li et al., 2023; Wang et al., 2023; Sundermeyer et al., 2021; Morrison et al., 2018; Enwerem et al., 2026a;c). ParcelStow evaluates grasps acquired during task execution and uses the force closure margin (Murray et al., 2017) only to characterize realized hand–object contacts.

1.1.3 Policy Evaluation and Controlled Perturbations

Robot learning benchmarks specify tasks, demonstrations, and controlled tests of generalization (Yu et al., 2019; James et al., 2019; Mees et al., 2021; Liu et al., 2023; Zhu et al., 2020; Gu et al., 2023; Tao et al., 2024; Nasiriany et al., 2024). Recent work also studies evaluation procedures. Li et al. (2025) compare relative policy performance in simulation and hardware, Zhou et al. (2025) automate repeated hardware trials, Pumacay et al. (2024) evaluate task success under fourteen environmental perturbations, Sagar et al. (2025) search structured environment variations, and Jiang et al. (2026) audit benchmark scores for shortcut solutions, statistical significance, benchmark overfitting, and dependence on demonstration sources. These studies examine environmental variation and the reliability of policy evaluation. We evaluate the learner and its demonstration expert under a common phase-duration law, with scene geometry and task goals fixed. We report their success curves and signed success-rate difference across speedup factors.

2 Problem Formulation

We consider a manipulation task with physical state $x_t \in \mathcal{X}$ at control step t, action $u_t \in \mathcal{U}$, and K ordered task phases $\mathcal{K} = \{k_1, \ldots, k_K\}$. The initial condition $\omega \in \Omega$ specifies the randomized components of the initial physical state through $x_0 = x_{\text{init}}(\omega)$. Let P_Ω be their distribution over the initial-condition space Ω. The task also maintains a normalized phase variable $\phi_t \in [0, 1]$, the previous action u_{t-1}, and the speedup factor r defining the phase durations in Eq. (5). We collect these variables in the augmented state $z_t = (x_t, \phi_t, u_{t-1}, r) \in \mathcal{Z}$, whose controlled transition satisfies $z_{t+1} = F(z_t, u_t)$. The learner observes $o_t = g(z_t)$ through $g : \mathcal{Z} \to \mathcal{O}$ and acts according to $u_t = \pi_L(o_{0:t})$. The expert acts according to $u_t = \pi_E(h_E(z_{0:t}))$, where h_E maps the augmented state history to expert inputs. This notation allows each policy to maintain internal state, which resets at episode boundaries. Both policies act through the same action space, physical dynamics, task geometry, and success criterion. The speedup factor lies in an evaluation set $\mathcal{R}_E \subset \mathbb{R}_{>0}$, and $r_0 = 1$ denotes nominal execution. Demonstration collection samples speedup factors from P_D, supported on the demonstrated speed range $\mathcal{R}_D \subset \mathbb{R}_{>0}$.

The task schedule at speedup factor r determines a horizon of $H(r)$ policy steps. For policy π, speedup factor r, and initial condition ω, let

$$\tau(\pi, r, \omega) = \big(z_0, u_0, \ldots, z_{H(r)}\big) \tag{1}$$

denote the resulting rollout. The binary function $Y(\tau) \in \{0, 1\}$ equals one if and only if the rollout satisfies the task success criterion. The policy's task success probability at r is

$$p_\pi(r) = \Pr[Y(\tau(\pi, r, \omega)) = 1]\,, \tag{2}$$

where the probability is over $\omega \sim P_\Omega$ and any internal randomness of π. Given N evaluation draws $\omega_1, \ldots, \omega_N$, we estimate $p_\pi(r)$ as

$$\widehat{p}_\pi(r) = \frac{1}{N} \sum_{i=1}^{N} Y(\tau(\pi, r, \omega_i))\,. \tag{3}$$

The learner is trained from a set of M expert demonstrations

$$\mathcal{D}_E = \{\tau_m\}_{m=1}^{M}\,. \tag{4}$$

Expert rollouts draw speedup factors and randomized initial conditions from P_D and P_Ω, respectively; successful completed rollouts form $\mathcal{D}_E$. DAgger augments this initial set with expert labels at learner-visited states. We then evaluate learner performance relative to the expert as r varies. We focus on learners that match or exceed the expert's observed nominal success. Evaluations of other learners characterize performance from lower nominal success levels or failure when nominal success is near zero.

3 METHODOLOGY

We estimate expert and learner success rates at each $r \in \mathcal{R}_E$. The resulting success-versus-speed curves characterize each policy's response to the timing perturbation.

3.1 EXECUTION-SPEED PERTURBATION

Let $\mathcal{M} \subseteq \mathcal{K}$ denote the task phases whose durations vary. If phase k has nominal duration T_k, we set its duration at speedup factor r to

$$T_k(r) = \begin{cases} T_k/r, & k \in \mathcal{M}, \\ T_k, & k \notin \mathcal{M}. \end{cases} \tag{5}$$

During phase k_j, the phase variable is $\phi_t = (j - 1 + f_t)/K$, where $f_t \in [0, 1]$ is the elapsed fraction of that phase. Changing r changes the phase schedule and the number of policy steps $H(r)$ while leaving the task geometry, physical parameters, and success criterion fixed. The learner receives both ϕ_t and r in its observation, so the execution-speed intervention changes the schedule and the speed input together. We hold each policy's action horizon and replanning frequency fixed across speedup factors. Task-specific phase choices and demonstrated ranges instantiate this intervention in Sec. 4.1.3.

3.2 EXPERT–LEARNER COMPARISON

We pair expert and learner rollouts at each speedup factor r using shared initial-condition draws; Sec. 4.2 specifies the verification available for each task. Their signed empirical difference is

$$G_L(r) = \widehat{p}_{\pi_E}(r) - \widehat{p}_{\pi_L}(r). \tag{6}$$

Positive values indicate higher expert success on the matched evaluation sample. We examine two cases: exact observed nominal equality, $G_L(1) = 0$, and higher observed nominal learner success, $G_L(1) < 0$. Neither condition is a test of population equivalence or superiority. We compare the observed nominal ordering with the ordering at other speedup factors, distinguishing factors inside $\mathcal{R}_D$ from extrapolation outside it. Reporting both success curves alongside $G_L(r)$ distinguishes agreement at high success from agreement when both policies fail. Near-zero nominal success restricts a policy to failure characterization.

Policy comparisons require a common task, action interface, success criterion, phase-duration law, and initial-condition distribution. For conditions without verified pairing or distribution matching, we report only standalone success rates. We use pointwise Wilson intervals for success rates and bootstrap complete episode pairs at a common speedup factor for uncertainty in $G_L(r)$. The task-specific evaluation conditions are given in Sec. 4.2.

3.3 FAILURE LOCALIZATION

We localize performance differences by comparing task-stage outcomes and terminal failure categories across policies and speedup factors. Stage outcomes record which physical events occur during a rollout; terminal categories distinguish how unsuccessful rollouts end.

4 EXPERIMENTS

4.1 TASKS AND POLICY INTERFACE

4.1.1 SIMULATION AND PHYSICAL TASKS

We apply the evaluation procedure in Sec. 3 to three manipulation tasks in Isaac Lab (Mittal et al., 2023), using a fixed-base Unitree G1 and a RealHand L6 right hand. The objects are dynamically free rigid bodies with mass 0.12 kg and fixed nominal friction $\mu = 0.5$. In the PARCEL task, the robot must acquire an $80 \times 55 \times 40$ mm cuboid, reorient it by $90°$, and insert it into an open-front receptacle. In UPRIGHT, the robot is tasked with retrieving and placing a $55 \times 55 \times 180$ mm object at a target region in a vertical orientation. In PEG, the robot is tasked with inserting a cuboidal object

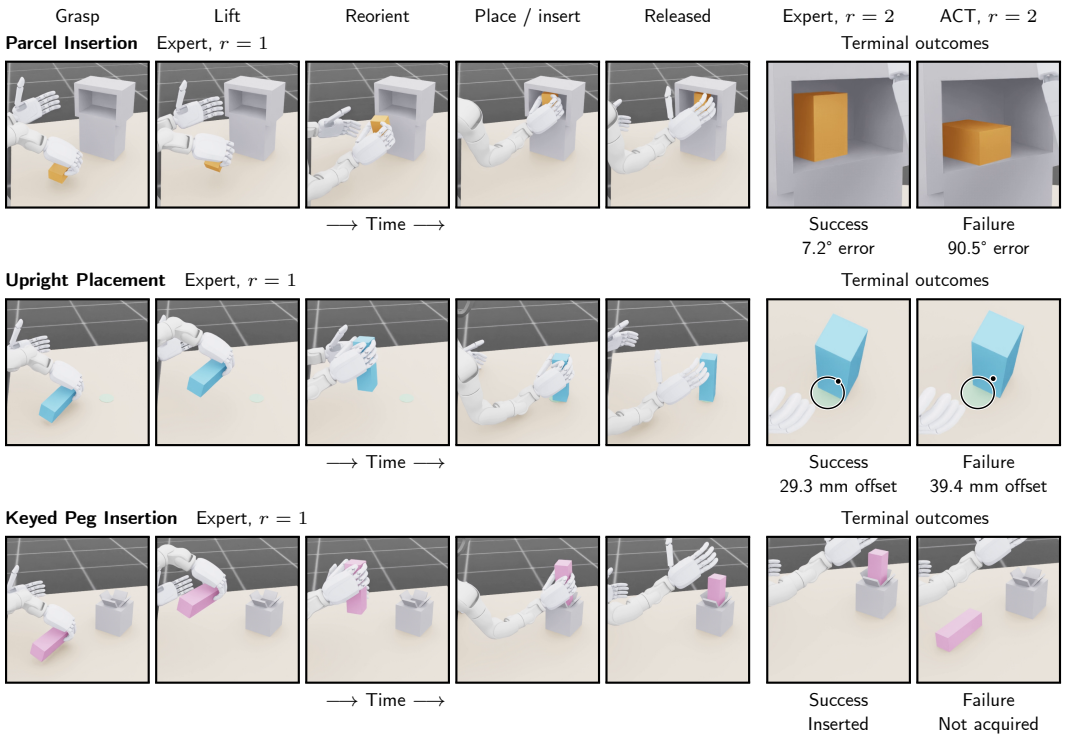


Figure 2: **Recorded Task Executions and Failure Examples.** The first five columns show nominal expert execution ($r = 1$). The last two columns show enlarged expert and ACT terminal states at $r = 2$. PARCEL labels give orientation error (10° tolerance). The blue outline in the UPRIGHT column marks the projected 30 mm target boundary, with the dot marking the projected object-base center. For both UPRIGHT and PEG, $r = 2$ halves the scaled phase durations and lies outside the demonstration range.

(the *key* or *peg*) of compatible dimensions into a hollow box through a tight pocket. Fig. 2 shows recorded expert rollouts and ACT terminal failures for the three tasks.

Monitors record physical events and assign terminal failure categories. The monitor stores each event's occurrence until episode reset. Terminal PARCEL success requires insertion, release, settling, and final containment with orientation error at most 10°. UPRIGHT success requires placement, release, settling, and a final object base center within the 30 mm target radius with tilt at most 5°. PEG success requires containment at sufficient insertion depth, release, settling, and final tilt at most 5°. Intermediate events need not occur in sequence: insertion in PEG, for example, does not require prior held alignment. Event thresholds, downstream dependencies, and failure categories appear in Appendices A.1, A.2 and D.1.

4.1.2 OBSERVATION, ACTION, AND EXECUTION SPEED

All learners receive a 147-dimensional state observation containing full-body joint position and velocity, previous action, object pose, fingertip positions and contact forces, task phase, and the raw speedup factor r. Their 16-dimensional actions specify normalized joint-position offsets for the waist, right arm, and six actuated hand joints:

$$q_{\text{target},t} = q_{\text{default}} + 0.5u_t. \tag{7}$$

Here, $q_{\text{target},t}, q_{\text{default}} \in \mathbb{R}^{16}$ are the target and default joint-angle vectors in radians, and 0.5 is the angle scale per unit action. Implicit proportional–derivative (PD) joint drives track the targets at a 50 Hz control interface; gains and simulation settings appear in Appendix A.4. The experts interpolate task-specific joint-space knots solved offline by inverse kinematics and apply runtime correction through the same action channel (Appendix A.4.1). They do not use the learner observation vector.

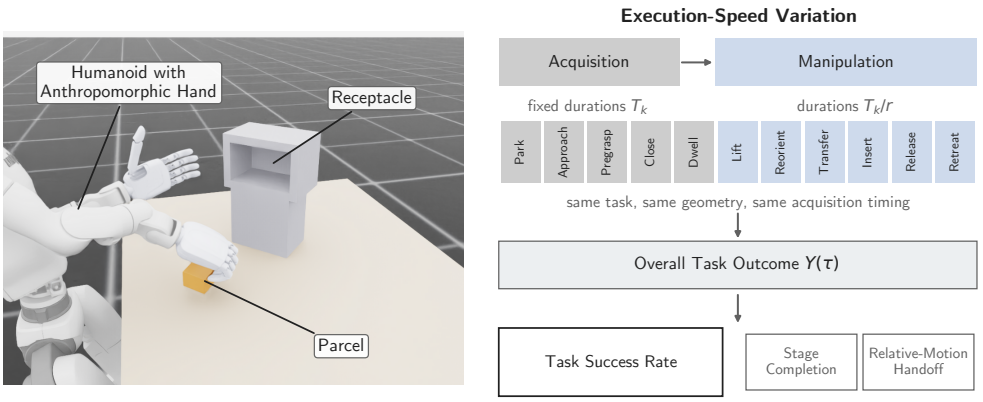


Figure 3: **Execution-Timing Perturbation Illustrated on PARCEL.** All three tasks keep acquisition timing fixed and divide selected manipulation-phase durations by the speedup factor r. The parcel scene and phase sequence illustrate this rule; T_k denotes a nominal phase duration. All tasks yield success and stage outcomes. The handoff diagnostic is evaluated only on PARCEL (Sec. 4.3).

Table 1: Task-Specific Demonstration and Evaluation Settings. Here, M is the number of demonstrations. N is final episodes per actor and speedup factor. UPRIGHT/PEG ACT use 80/20 episode partitions, while their DP and DAgger use all admitted demonstrations. The listed random seeds initialize the generators of initial-condition samples. Mint marks demonstration support containing the highlighted higher-factor comparison, and salmon marks support excluding it.

Task	M	$\mathcal{R}_D$	N	ACT selection	Final bank
PARCEL	297	$[0.5, 2]$	100	Fixed epoch 2000	Seed 12345
UPRIGHT	315	$[0.75, 1.75]$	200	100 episodes, seed 42001	Seed 73001
PEG	325	$[0.5, 1]$	200	100 episodes, seed 42001	Seed 73001

4.1.3 SPEEDUP FACTORS AND DEMONSTRATIONS

Acquisition phases retain fixed durations. Speedup factor r scales the later lift, reorientation, transport, placement or insertion, release, and retreat phases according to Eq. (5), while the settling window stays fixed. Fig. 3 illustrates this common timing rule using PARCEL. The expert's geometric reference path and task tolerances do not change. For PARCEL, the total scheduled cycle is $6.3 + 7.8/r$ seconds. The task-specific demonstrated ranges and evaluation banks are summarized in Table 1. Speedup factors outside the demonstrated interval are extrapolation, including UPRIGHT $r = 2$ and PEG $r = 1.5$.

4.1.4 TRAINING AND CHECKPOINT SELECTION

Resets use 10 mm planar object-position jitter and joint-position perturbations, with fixed mass and friction; Appendix A.3 specifies P_Ω. On PARCEL, ACT-A/B/C denote three ACT training runs; their random seeds are listed in Table 7.

UPRIGHT and PEG use ACT training with reconstruction loss averaged over valid action elements, episode-level train/validation partitions, and normalization computed on training episodes only. A 2,000-epoch budget retains validation-ranked candidates. We select a checkpoint using 100 nominal development episodes, ordered by success, task-stage counts, validation loss, and earliest epoch. Before final evaluation, we designated one trained ACT policy per task for the comparisons in Fig. 1 and two additional training runs for replications. Final outcomes do not select policies or checkpoints. Appendices B.2 and B.3 give the complete procedure.

Diffusion Policy and DAgger use the final training epoch and final aggregation round, respectively, without selection on nominal success. DP predicts a horizon of 16 from two observation steps and

executes eight actions. DAgger trains an MLP on demonstrations and expert labels at learner-visited states. Architecture and optimization settings are given in Appendix B.

4.2 Evaluation and Statistical Treatment

PARCEL uses speedup factors 0.5, 1, 1.5, 2, 2.25, 2.5, and 3. UPRIGHT and PEG use speedup factors 0.5, 0.75, 1, 1.25, 1.5, 1.75, 2, and 2.5. Final evaluations disable observation corruption and action noise. For the expert–ACT comparisons in Figs. 1 and 6, episode pairs match on episode index, recorded initial object pose, and evaluation settings. We use common random seeds to generate their initial states; a seed initializes the pseudorandom number generator. UPRIGHT and PEG share indexed initial-condition samples. PARCEL initial joint states were not recorded, so pairing assumes that common seeds and reset settings reproduce those states.

We report counts and pointwise 95% Wilson intervals. For same-factor pairs established on this basis, 20,000 bootstrap resamples of complete episode pairs give the central 95% interval for $G_L(r)$. These intervals do not establish population equivalence or account for training-seed variability. Different speedup factors use different seeds and are not paired across speedup factors. Reproducibility settings are collected in Table 7.

PARCEL DP differs from the expert in recorded initial pose on 476/700 episode keys. PARCEL DAgger differs on 80/700, at $r = 2$ and $r = 2.5$. Full reset-distribution matching is not established for these conditions, so they receive standalone rates without expert-difference inference. UPRIGHT/PEG DP and DAgger are paired with the expert at each speedup factor.

4.3 Parcel Failure Diagnostics

We compare PARCEL stage outcomes and terminal failures. To assess retention independently of insertion success, we replace each actor's downstream commands after acquisition with a shared controller that applies the expert's relative hand-pose commands, anchored at that actor's command-derived handoff pose. Retention is evaluated until insertion begins or the parcel contacts the receptacle. We also compute hand–parcel relative motion, arm joint velocity utilization, and the Ferrari–Canny margin of realized contacts (Ferrari & Canny, 1992). Appendix D specifies the handoff trigger, controller, endpoints, and contact analysis.

5 Results and Discussion

5.1 Equal Nominal Success Does Not Persist Within Demonstrated Parcel Speedup Factors

Fig. 1a compares the PARCEL expert with ACT-A. The expert and ACT-A each succeed on 100/100 nominal episodes. At $r = 1.5$ they succeed on 99/100 and 91/100, and at the maximum demonstrated speedup factor $r = 2$ on 84/100 and 53/100. The expert-minus-ACT difference at $r = 2$ is 31 percentage points with paired bootstrap 95% interval $[18, 44]$ points. This separation occurs within the demonstrated speed range.

The paired PARCEL outcomes at $r = 2$ are 42 successes for both actors, 42 for the expert alone, 11 for ACT-A alone, and five failures for both. Thus ACT-A's failure set is not a superset of the expert's. At $r = 0.5$, ACT-A also trails the expert by 16 points, so exact nominal equality does not persist at either side of nominal speed. At $r = 3$, both fail all 100 episodes; equality at the floor is not performance preservation. Full paired curves, seed replications, and demonstration-subset results appear in Appendices C.2 and C.3.

5.2 Higher Nominal ACT Success Reverses Beyond Demonstrated Speedup Factors

On UPRIGHT, nominal ACT success is 194/200 versus expert 185/200. At $r = 2$, ACT succeeds on 24/200 and the expert on 105/200, an expert-minus-ACT gap of 40.5 points with paired 95% interval $[32, 49]$ points. On PEG, nominal ACT success is 191/200 versus expert 182/200. At $r = 1.5$, the

Table 2: Success Counts for Three Imitation Methods and Their Experts. Nominal and selected higher-factor conditions are shown. PARCEL $r = 2$ is demonstrated, while UPRIGHT $r = 2$ and PEG $r = 1.5$ are extrapolation. PARCEL DP and the displayed higher-factor DAgger cell are standalone estimates because initial-condition pairing fails. DAgger is reported as failure characterization, not speed degradation. Mint marks nominal expert and ACT counts; salmon marks ACT at the higher speedup factor.

Task	r	Expert	ACT	Diffusion Policy	DAgger
PARCEL	1	100/100	100/100	68/100	3/100
PARCEL	2	84/100	53/100	54/100	3/100
UPRIGHT	1	185/200	194/200	181/200	1/200
UPRIGHT	2	105/200	24/200	57/200	2/200
PEG	1	182/200	191/200	95/200	0/200
PEG	1.5	150/200	3/200	25/200	0/200

counts are 3/200 and 150/200, a gap of 73.5 points with interval $[67.5, 79.5]$ points. The observed nominal ACT advantage is 4.5 points on each task.

The distinction from PARCEL is the demonstration support: UPRIGHT $r = 2$ exceeds 1.75, and PEG $r = 1.5$ exceeds 1.0. These rankings therefore reverse under extrapolation. Response to the speedup factor is also not uniformly monotonic. UPRIGHT ACT succeeds on 199/200 at $r = 1.25$ and 194/200 at $r = 1.75$ before its larger decrease. PEG ACT succeeds on 166/200 and 171/200 at $r = 0.5$ and 0.75, below its nominal 191/200. The full curves therefore contain information absent from the nominal ranking.

The task-specific failure categories differ. For the UPRIGHT task at $r = 2$, both expert and ACT acquire in 200/200 episodes, but the post-release failure category accounts for 91 expert and 173 ACT failures. For the PEG task at $r = 1.5$, ACT acquires in only 141/200, with 59 acquisition failures and 133 transport losses, although acquisition timing is fixed. The policy receives the speedup factor during acquisition (Sec. 4.1.2). These observations do not isolate whether speedup-factor conditioning, action prediction, or later state-distribution changes cause the failures.

5.3 DIFFUSION POLICY AND DAGGER RESULTS

Table 2 reports ACT, DP, and DAgger alongside each task expert. DP nominal success is 68/100 on PARCEL, 181/200 on UPRIGHT, and 95/200 on PEG. Their curves characterize speed response from these task-specific nominal success levels. For the PEG task at $r = 1.25$, DP reaches 120/200, exceeding its own nominal count before falling to 25/200 at $r = 1.5$.

DAgger nominal successes are 3/100 on PARCEL, 1/200 on UPRIGHT, and 0/200 on PEG. Across all speedup factors, its maxima are 3/100, 4/200, and 1/200, respectively. These floor-saturated curves support failure characterization, not execution-speed degradation. Complete success counts and ACT replications appear in Appendix C, with checkpoint-selection details in Appendix B.3.

5.4 PARCEL DIAGNOSTICS LOCALIZE FAILURES TO INSERTION

For the PARCEL task at $r = 2$, ACT-A acquires and lifts in all 100 episodes, records 89 insertions, and completes 53 episodes. Its failures comprise 35 insertion misalignments, 10 jams, one excessive-slip category, and one timeout. The expert reaches the pre-insertion predicate in all 100 episodes and records 16 insertion jams. Stage counts and terminal categories localize these failures to insertion and final orientation (Appendix D.1.1).

In the handoff experiment defined in Sec. 4.3, expert and ACT-A each retain all 300 acquired parcels to the free-space endpoint, defined by insertion onset or first receptacle contact, over $r \in \{1, 1.5, 2\}$. Continuing the shared relative-motion controller through insertion at $r = 2$ yields 95/100 expert and 64/100 ACT-A successes. The intervention preserves each actor's realized handoff state, so it cannot distinguish state error from subsequent contact effects. The diagnostic results are reported in Appendices D.1.1 and D.2.1.

5.5 Variation Across Training Seeds

The magnitude of the expert–learner success-rate difference depends on the trained policy. Across three ACT training runs, nominal PARCEL success ranges from 62% to 100%, and UPRIGHT success at $r = 2$ ranges from 24/200 to 98/200, compared with expert 105/200 (Tables 9 and 11). Each curve represents one trained policy; its pointwise intervals quantify variation across evaluation episodes (Sec. 4.2). The results concern state-based policies evaluated under execution-timing perturbations in simulation, with task geometry and friction fixed.

6 Conclusion

We evaluated expert–learner performance variation under a controlled perturbation of a task condition. ParcelStow varies designated phase durations through the speedup factor in Eq. (5), with task geometry, physical parameters, and success criteria fixed. Across the three tasks, we observe two forms of performance separation: on PARCEL, expert and ACT have equal observed nominal success but differ by 31 percentage points at the maximum demonstrated speedup factor; on UPRIGHT and PEG, higher observed nominal ACT success reverses beyond the respective demonstration ranges. DP provides additional success curves; DAgger's near-zero nominal success permits only failure characterization. Nominal task success alone does not establish preservation of expert performance under task-condition perturbations; the timing experiments provide a counterexample.

AI Use Statement

All core research tasks—including conceptual design, experimental procedures, and data analysis—were performed by the authors. Generative AI tools were used solely for targeted technical assistance in code optimization, literature scoping, and editing select portions of the manuscript's text for clarity. The authors rigorously verified all AI-assisted code and text, and retain full accountability for the accuracy and final content of this work.

References

Kevin Black, Noah Brown, Danny Driess, et al. π_0: A Vision-Language-Action Flow Model for General Robot Control, 2024.

Kevin Black, Manuel Y. Galliker, and Sergey Levine. Real-Time Execution of Action Chunking Flow Policies. In *Advances in Neural Information Processing Systems (NeurIPS)*, 2025. arXiv:2506.07339.

Cheng Chi, Siyuan Feng, Yilun Du, Zhenjia Xu, Eric Cousineau, Benjamin Burchfiel, and Shuran Song. Diffusion Policy: Visuomotor Policy Learning via Action Diffusion. In *Robotics: Science and Systems (RSS)*, 2023. doi: 10.15607/RSS.2023.XIX.026.

Clinton Enwerem, John S. Baras, and Calin Belta. FIRMGrasp: A Friction-Informed Risk Margin for Robust Grasp Synthesis, 2026a. arXiv:2607.25049.

Clinton Enwerem, John S. Baras, and Calin Belta. EquiDexFlow: Contact-Grounded SE(3)-Equivariant Dexterous Grasp Generative Flows, 2026b. arXiv:2606.12728.

Clinton Enwerem, Shreya Kalyanaraman, John S. Baras, and Calin Belta. Variational Neural Belief Parameterizations for Robust Dexterous Grasping under Multimodal Uncertainty, 2026c. arXiv:2604.25897.

Carlo Ferrari and John F. Canny. Planning Optimal Grasps. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 2290–2295, 1992. doi: 10.1109/ROBOT.1992.219918.

Jiayuan Gu, Fanbo Xiang, Xuanlin Li, et al. ManiSkill2: A Unified Benchmark for Generalizable Manipulation Skills, 2023.

Lingxiao Guo, Zhengrong Xue, Zijing Xu, and Huazhe Xu. DemoSpeedup: Accelerating Visuomotor Policies via Entropy-Guided Demonstration Acceleration, 2025.

Harsh Gupta, Mohammad Amin Mirzaee, and Wenzhen Yuan. Grasp to Act: Dexterous Grasping for Tool Use in Dynamic Settings, 2026.

Stephen James, Zicong Ma, David Rovick Arrojo, and Andrew J. Davison. RLBench: The Robot Learning Benchmark & Learning Environment, 2019.

Xiaogang Jia, Atalay Donat, Xi Huang, Xuan Zhao, Denis Blessing, Hongyi Zhou, Han A. Wang, Hanyi Zhang, Qian Wang, Rudolf Lioutikov, and Gerhard Neumann. X-IL: Exploring the Design Space of Imitation Learning Policies, 2025.

Tianchong Jiang, Xiangshan Tan, Samuel Wheeler, Luzhe Sun, Tewodros W. Ayalew, and Matthew Walter. What Are We Actually Benchmarking in Robot Manipulation?, 2026.

Kushal Kedia, Prithwish Dan, Angela Chao, Maximus Adrian Pace, and Sanjiban Choudhury. One-Shot Imitation under Mismatched Execution, 2024.

Albert H. Li, Preston Culbertson, Joel W. Burdick, and Aaron D. Ames. FRoGGeR: Fast Robust Grasp Generation via the Min-Weight Metric. In *2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 6809–6816, 2023. doi: 10.1109/IROS55552.2023.10341806. arXiv:2302.13687.

Xuanlin Li, Kyle Hsu, Jiayuan Gu, Oier Mees, Karl Pertsch, Homer Rich Walke, Chuyuan Fu, Ishikaa Lunawat, Isabel Sieh, Sean Kirmani, Sergey Levine, Jiajun Wu, Chelsea Finn, Hao Su, Quan Vuong, and Ted Xiao. Evaluating Real-World Robot Manipulation Policies in Simulation. In *Proceedings of the 8th Conference on Robot Learning (CoRL)*, volume 270 of *Proceedings of Machine Learning Research*, pp. 3705–3728, 2025.

Haoran Liao, Pengyue Wang, Shuoyu Chen, Kehan Cheng, Xuhang Chen, Yuhao Lin, Mu Lin, Zhizhao Liang, Xiaoyi Fan, Chengyi Xing, Dan Niu, Yi-Lin Wei, and Wei-Shi Zheng. Dynamic-Manip: Enabling Dynamic Manipulation from a Single Static Demonstration, 2026.

Bo Liu, Yifeng Zhu, Chongkai Gao, Yihao Feng, Qiang Liu, Yuke Zhu, and Peter Stone. LIBERO: Benchmarking Knowledge Transfer for Lifelong Robot Learning, 2023.

Ajay Mandlekar, Danfei Xu, Josiah Wong, Soroush Nasiriany, Chen Wang, Rohun Kulkarni, Li Fei-Fei, Silvio Savarese, Yuke Zhu, and Roberto Martín-Martín. What Matters in Learning from Offline Human Demonstrations for Robot Manipulation, 2021.

Oier Mees, Lukas Hermann, Erick Rosete-Beas, and Wolfram Burgard. CALVIN: A Benchmark for Language-Conditioned Policy Learning for Long-Horizon Robot Manipulation Tasks, 2021.

Mayank Mittal, Calvin Yu, Qinxi Yu, Jingzhou Liu, Nikita Rudin, David Hoeller, Jia Lin Yuan, Ritvik Singh, Yunrong Guo, Hammad Mazhar, Ajay Mandlekar, Buck Babich, Gavriel State, Marco Hutter, and Animesh Garg. Orbit: A Unified Simulation Framework for Interactive Robot Learning Environments. *IEEE Robotics and Automation Letters*, 8(6):3740–3747, 2023. doi: 10.1109/LRA.2023.3270034.

Douglas Morrison, Juxi Leitner, and Peter Corke. Closing the Loop for Robotic Grasping: A Real-time, Generative Grasp Synthesis Approach. In *Robotics: Science and Systems (RSS)*, 2018. doi: 10.15607/RSS.2018.XIV.021.

Richard M. Murray, Zexiang Li, and S. Shankar Sastry. *A Mathematical Introduction to Robotic Manipulation*. CRC Press, 1 edition, 2017. ISBN 978-1-315-13637-0.

Taewook Nam, Junmo Cho, Youngsoo Jang, and Sung Ju Hwang. SpeedAug: Policy Acceleration via Tempo-Enriched Policy and RL Fine-Tuning, 2026.

Soroush Nasiriany, Abhiram Maddukuri, Lance Zhang, Adeet Parikh, Aaron Lo, Abhishek Joshi, Ajay Mandlekar, and Yuke Zhu. RoboCasa: Large-Scale Simulation of Everyday Tasks for Generalist Robots, 2024.

Junnan Nie, Jiayi Li, Jiachen Zhang, Junyi Lao, Chenghao Liu, Tianle Zhang, Liang Lin, and Songfang Huang. PACE: Phase-Aware Chunk Execution for Robot Policies with Action Chunking, 2026.

Octo Model Team, Dibya Ghosh, et al. Octo: An Open-Source Generalist Robot Policy, 2024.

Open X-Embodiment Collaboration. Open X-Embodiment: Robotic Learning Datasets and RT-X Models, 2023.

OpenAI, Marcin Andrychowicz, Bowen Baker, et al. Learning Dexterous In-Hand Manipulation, 2018.

OpenAI, Ilge Akkaya, Marcin Andrychowicz, et al. Solving Rubik's Cube with a Robot Hand, 2019.

Aleksei Petrenko, Arthur Allshire, Gavriel State, Ankur Handa, and Viktor Makoviychuk. DexPBT: Scaling up Dexterous Manipulation for Hand-Arm Systems with Population Based Training. In *Robotics: Science and Systems (RSS)*, 2023. doi: 10.15607/RSS.2023.XIX.037.

Wilbert Pumacay, Ishika Singh, Jiafei Duan, Ranjay Krishna, Jesse Thomason, and Dieter Fox. THE COLOSSEUM: A Benchmark for Evaluating Generalization for Robotic Manipulation. In *Robotics: Science and Systems (RSS)*, 2024. arXiv:2402.08191.

Aravind Rajeswaran, Vikash Kumar, Abhishek Gupta, Giulia Vezzani, John Schulman, Emanuel Todorov, and Sergey Levine. Learning Complex Dexterous Manipulation with Deep Reinforcement Learning and Demonstrations, 2017.

Nadun Ranawaka Arachchige, Zhenyang Chen, Wonsuhk Jung, Woo Chul Shin, Rohan Bansal, Pierre Barroso, Yu Hang He, Yingyan Celine Lin, Benjamin Joffe, Shreyas Kousik, and Danfei Xu. SAIL: Faster-than-Demonstration Execution of Imitation Learning Policies. In *Proceedings of the 9th Conference on Robot Learning (CoRL)*, volume 305 of *Proceedings of Machine Learning Research*, pp. 721–749, 2025.

Máximo A. Roa and Raúl Suárez. Grasp Quality Measures: Review and Performance. *Autonomous Robots*, 38(1):65–88, 2015. ISSN 1573-7527. doi: 10.1007/s10514-014-9402-3.

Stéphane Ross, Geoffrey J. Gordon, and J. Andrew Bagnell. A Reduction of Imitation Learning and Structured Prediction to No-Regret Online Learning. In *Proceedings of the 14th International Conference on Artificial Intelligence and Statistics (AISTATS)*, volume 15 of *Proceedings of Machine Learning Research*, pp. 627–635, 2011.

Som Sagar, Jiafei Duan, Sreevishakh Vasudevan, Yifan Zhou, Heni Ben Amor, Dieter Fox, and Ransalu Senanayake. From Mystery to Mastery: Failure Diagnosis for Improving Manipulation Policies, 2025. URL `https://arxiv.org/abs/2412.02818v2`. Version 2.

Nur Muhammad Mahi Shafiullah, Zichen Jeff Cui, Ariuntuya Altanzaya, and Lerrel Pinto. Behavior Transformers: Cloning k Modes with One Stone, 2022.

Mohit Shridhar, Lucas Manuelli, and Dieter Fox. Perceiver-Actor: A Multi-Task Transformer for Robotic Manipulation, 2022.

Jean-Pierre Sleiman, Mayank Mittal, and Marco Hutter. Guided Reinforcement Learning for Robust Multi-Contact Loco-Manipulation. In *Proceedings of the 8th Conference on Robot Learning (CoRL)*, volume 270 of *Proceedings of Machine Learning Research*, pp. 531–546, 2025.

Martin Sundermeyer, Arsalan Mousavian, Rudolph Triebel, and Dieter Fox. Contact-GraspNet: Efficient 6-DoF Grasp Generation in Cluttered Scenes. In *IEEE International Conference on Robotics and Automation (ICRA)*, pp. 13438–13444, 2021. doi: 10.1109/ICRA48506.2021.9561877.

Stone Tao, Fanbo Xiang, Arth Shukla, Yuzhe Qin, Xander Hinrichsen, Xiaodi Yuan, Chen Bao, Xinsong Lin, Yulin Liu, Tse-kai Chan, Yuan Gao, Xuanlin Li, Tongzhou Mu, Nan Xiao, Arnav Gurha, Zhiao Huang, Roberto Calandra, Rui Chen, Shan Luo, and Hao Su. ManiSkill3: GPU Parallelized Robotics Simulation and Rendering for Generalizable Embodied AI, 2024.

Vitalis Vosylius and Edward Johns. Instant Policy: In-Context Imitation Learning via Graph Diffusion. In *International Conference on Learning Representations (ICLR)*, 2025. arXiv:2411.12633.

Ruicheng Wang, Jialiang Zhang, Jiayi Chen, Yinzhen Xu, Puhao Li, Tengyu Liu, and He Wang. DexGraspNet: A Large-Scale Robotic Dexterous Grasp Dataset for General Objects Based on Simulation. In *2023 IEEE International Conference on Robotics and Automation (ICRA)*, pp. 11359–11366. IEEE, 2023. doi: 10.1109/ICRA48891.2023.10160982. arXiv:2210.02697.

Tianhe Yu, Deirdre Quillen, Zhanpeng He, et al. Meta-World: A Benchmark and Evaluation for Multi-Task and Meta Reinforcement Learning, 2019.

Hui Zhang, Zijian Wu, Linyi Huang, Sammy Christen, and Jie Song. RobustDexGrasp: Robust Dexterous Grasping of General Objects from Single-view Perception, 2025.

Tony Z. Zhao, Vikash Kumar, Sergey Levine, and Chelsea Finn. Learning Fine-Grained Bimanual Manipulation with Low-Cost Hardware. In *Robotics: Science and Systems (RSS)*, 2023. doi: 10.15607/RSS.2023.XIX.016.

Zhiyuan Zhou, Pranav Atreya, You Liang Tan, Karl Pertsch, and Sergey Levine. AutoEval: Autonomous Evaluation of Generalist Robot Manipulation Policies in the Real World. In *Proceedings of the 9th Conference on Robot Learning (CoRL)*, volume 305, pp. 1997–2017, 2025.

Yuke Zhu, Josiah Wong, Ajay Mandlekar, Roberto Martín-Martín, Abhishek Joshi, Kevin Lin, Abhiram Maddukuri, Soroush Nasiriany, and Yifeng Zhu. robosuite: A Modular Simulation Framework and Benchmark for Robot Learning, 2020.

APPENDIX

A TASK SPECIFICATIONS AND CONTROL

A.1 PARCEL PHYSICAL TASK SPECIFICATION

Table 3 summarizes the PARCEL task settings. We fixed these settings using kinematic probes and expert-only calibration (Appendix B.1), without using learner outcomes. Table 4 gives the success predicates evaluated on simulator state; none uses a wrench-space quantity.

Table 3: PARCEL Task and Evaluation Settings.

Quantity	Value
Parcel extents and mass	$80 \times 55 \times 40$ mm, 0.12 kg
Nominal friction	$\mu = 0.5$ static and dynamic
Start pose	$(0.35,\ 0,\ 0.721)$ m, yaw 45°
Reorientation	90° about the parcel width axis
Transport distance	0.334 m
Tight-axis insertion clearance	10 mm per side
Final orientation tolerance	10°
Policy and control frequency	50 Hz
Training speedup distribution	$r \sim \mathcal{U}[0.5, 2.0]$
Evaluation speedup grid	$\{0.5, 1, 1.5, 2, 2.25, 2.5, 3\}$
Fixed-phase duration	6.3 s
Duration of phases in $\mathcal{M}$	$7.8/r$ s

Table 4: Success Predicates of the Overall Task. Insertion and settling flags record events during the episode; task success additionally requires the final position and orientation tolerances.

Predicate	Condition
Acquisition	Parcel center rises by at least 20 mm while the thumb distal phalanx and at least one further distal phalanx each exert more than 1 N of contact force on the parcel.
Lift clearance	While held, parcel center rises at least 60 mm above the table.
Reorientation	While held before the insertion phase, parcel orientation error falls below 15°.
Pre-insertion	While held, position error from the pre-insertion target is below 30 mm and orientation error is below 15°.
Insertion	Parcel center passes at least 50 mm beyond the receptacle entrance plane with the parcel inside the receptacle cross section.
Release	After the insertion event, contact force between the distal phalanges and parcel falls below 0.5 N for at least 0.1 s.
Settling	After release, the insertion predicate holds for 0.4 s with linear speed below 0.02 m/s, angular speed below 0.2 rad/s. Orientation is tested separately at termination.
Terminal success	Insertion, release, and settling flags hold, and the final parcel is inside the receptacle with orientation error at most 10°.

A.2 UPRIGHT AND PEG SPECIFICATIONS

The UPRIGHT and PEG tasks use $55 \times 55 \times 180$ mm rigid objects of mass 0.12 kg and friction 0.5. UPRIGHT evaluates the base center within a 30 mm target radius and final tilt at most 5°, after placement, release, and settling. PEG uses 3 mm clearance, a 60 mm-deep pocket, an insertion depth threshold of 40 mm, and a final tilt tolerance of 5°. Acquisition predicates require opposed fingertip contact and a 20 mm rise. UPRIGHT phases after acquisition scale a 14.8 s manipulation interval with r and retain 6.7 s of fixed phases. PEG scales 15.8 s and retains 7.7 s. These schedules are task-specific, so the same dimensionless speedup factor does not imply equal physical velocities across tasks.

PEG event flags persist until reset, with the following dependencies. Lift, held-upright orientation, and held alignment require acquisition. Release requires insertion, settling requires release, and

terminal success requires insertion, release, settling, containment, and final tilt. Insertion requires neither prior alignment nor acquisition. UPRIGHT settling tests standing within the target with low linear and angular speed. PARCEL settling and final orientation are distinct tests (Table 4).

Table 5 specifies the intermediate events. In both tasks, acquisition requires a rise of at least 20 mm from the initial object center and contact forces above 1 N at the thumb distal phalanx and at least one other distal phalanx. For these predicates, held denotes acquisition without a recorded drop. In the release and drop predicates, contact force denotes the sum of distal-phalanx force magnitudes. Throughout settling in both tasks, linear speed is below 0.02 m/s and angular speed below 0.2 rad/s. Durations of 0.1, 0.2, and 0.4 s correspond to 5, 10, and 20 consecutive control steps. Scheduled phases advance with the task clock; event flags record physical predicates independently of that progression.

Table 5: Intermediate Event Predicates for UPRIGHT and PEG. Event flags persist until reset. Placement and held alignment use a 15-degree tilt tolerance; terminal success uses the stricter 5-degree tolerance.

Event	UPRIGHT	PEG
Lift clearance	While held, center rises at least 60 mm from its initial height.	While held, center rises at least 60 mm from its initial height.
Held upright	While held before release, tilt falls below 15 degrees.	While held before release, tilt falls below 15 degrees.
Placement / alignment	While held, center is within 30 mm of the placement target in 3D and tilt is below 15 degrees.	While held, planar center error from the pocket center is below 15 mm and tilt is below 15 degrees.
Insertion	Not applicable.	During insertion or a later phase, depth is at least 40 mm and all four base corners lie within the pocket cross section.
Release	After placement, contact force stays below 0.5 N for 0.1 s.	After insertion, contact force stays below 0.5 N for 0.1 s.
Settling	After release, base center stays within the 30 mm target radius and tilt below 15 degrees for 0.4 s.	After release, the insertion depth and containment predicates hold for 0.4 s.
Drop	While held before release, contact force is at most 0.5 N for 0.2 s and the placement predicate is currently false.	While held before release, contact force is at most 0.5 N for 0.2 s and the insertion predicate is currently false.

For PEG, insertion depth is the pocket-top height minus the object-center height plus the nominal object half-height. Pocket containment tests the transformed base corners. A jam is recorded when pocket contact exceeds 2 N for 0.2 s during the insertion phase without satisfying the insertion predicate.

Terminal categories use the first applicable condition after excluding task success. For UPRIGHT, the order is failure to acquire; a recorded transport drop; failure to place; failure to release (timeout); final tilt above 5 degrees or base center outside the target (post-release failure); and failure to settle (timeout). For PEG, the order is failure to acquire; a recorded transport drop; failure to insert (jam if recorded, otherwise alignment failure); failure to release or settle (timeout); final loss of containment; and final tilt above 5 degrees (alignment failure). All other cases are classified as other. Transport losses denote recorded drops before the release phase, whereas post-release failures denote the final pose category after placement and release.

A.3 INITIAL-CONDITION DISTRIBUTION

The distribution P_Ω randomizes object position and robot joint position. The two planar object offsets are independent uniform draws from $[-10, 10]$ mm about the task's nominal start; height and orientation take their nominal values. Joint offsets are independent uniform draws from $[-0.05, 0.05]$ rad about the default configuration, followed by clipping to the joint limits. Initial joint and object velocities are zero. UPRIGHT and PEG use indexed banks of these draws, shared across actors at each speedup factor. PARCEL uses the same reset distributions and seeds across the expert–ACT evaluations, with pairing verified for the recorded variables as specified in Sec. 4.2.

A.4 Observation, Action, and Control

The observation $o_t \in \mathcal{O} \subset \mathbb{R}^{147}$ contains full-body joint position and velocity (51 values each), the previous action u_{t-1} (16), parcel pose in the base frame (7), right-hand fingertip positions (15), right-hand fingertip contact forces (5), task phase (1), and speedup factor r (1). The 51 joint values cover every actuated and passively coupled degree of freedom in the G1 and L6 articulation; the fingertip terms cover five distal phalanges. The action $u_t \in \mathcal{U} \subset \mathbb{R}^{16}$ contains normalized joint position offsets for the waist (3), right arm and wrist (7), and six actuated right-hand joints (thumb roll and pitch and four finger MCP-pitch joints). The left arm, other hand joints, and passive coupling joints appear in the observation without corresponding action terms.

For the joint targets in Eq. (7), the proportional and derivative gains are K_p=5000, K_d=5 for the waist; K_p=300, K_d=10 for the arm; and K_p=10, K_d=0.2 for the hand. Simulation runs at 200 Hz (0.005 s per step), with four physics steps per control step, a fixed pelvis, 16 position solver iterations, and 8 velocity solver iterations.

A.4.1 Scripted Expert

The expert uses task phase, joint state, and controller correction state. Offline damped least-squares inverse kinematics converts task-space waypoints to waist and right-arm joint knots, using damping $\lambda = 0.05$, a null-space mid-range bias, a 0.15-rad step clamp, and tolerances of 4 mm and 2°. At runtime, the expert linearly interpolates these knots with eased acquisition transitions and applies integral correction ($k_i = 0.08$, clamped to ± 0.35) through the common joint-position interface. The knots, solver tolerances, and controller parameters are fixed across speedup factors; r changes only the scheduled phase durations in Eq. (5).

B Demonstrations, Training, and Evaluation

B.1 Parcel Speed Calibration

For expert speed calibration, we count successful episodes and compute arm joint velocity utilization, defined as

$$U_{\text{arm}} = \max_{t \geq t_{\text{lift}}} \max_{j \in \mathcal{J}_{\text{arm}}} |\dot{q}_j(t)| / \dot{q}_{j,\max}, \tag{8}$$

where t_{lift} is the first lift step, $\mathcal{J}_{\text{arm}}$ is the arm-joint set, $\dot{q}_j(t)$ is joint j's angular velocity, and $\dot{q}_{j,\max}$ is its velocity limit.

Expert-only calibration uses 64 episodes per candidate speedup factor with 1 cm planar start-position jitter (Table 6 and Fig. 4). Success is at least 58/64 through $r = 2$, then falls to 19/64 at $r = 2.5$ and 1/64 at $r = 3$. We therefore selected $[0.5, 2]$ as the demonstrated range. Arm joint velocity utilization stays below 0.3 across the candidate grid, so velocity saturation does not set this boundary. Calibration factors $r \in \{4, 5, 6\}$ lie outside the learner evaluation grid.

Table 6: Speed calibration using the expert, with 64 episodes per speedup factor, performed prior to learner evaluation. Mint denotes values with uniform success, while salmon denotes values with complete failure.

r	0.5	0.75	1.0	1.25	1.5	2.0	2.5	3.0	4.0	5.0	6.0
Success	63/64	63/64	63/64	64/64	64/64	58/64	19/64	1/64	0/64	0/64	3/64

B.2 ACT Architecture and Execution

All ACT instances use a chunk length of 100, a transformer encoder of 4 layers and a decoder of 7 layers, 8 attention heads, a model dimension of 512, a feedforward dimension of 3200, a 32-dimensional latent variable, and dropout 0.1. Training uses AdamW at learning rate 10^{-5} with weight decay 10^{-4}, batch size 8, and a KL weight of 10, for a budget of 2000 epochs. The Parcel instances train on the shared 297-episode set. Upright and Peg use the episode partitions and checkpoint selection in Appendix B.3. At each 50 Hz control step, ACT predicts a new chunk and

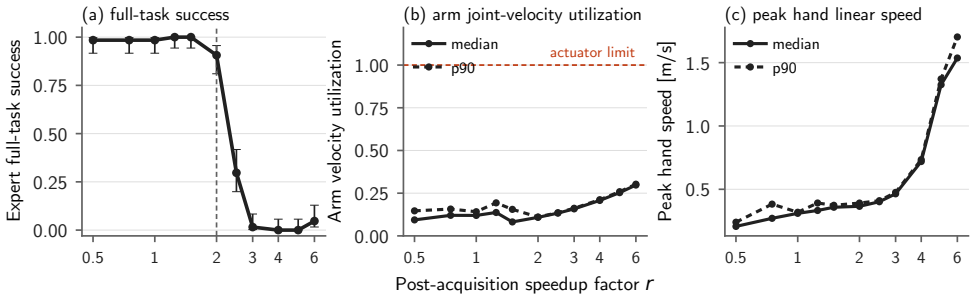


Figure 4: **Speed Calibration Using the Expert Over the Candidate Grid.** (a) Task success with Wilson 95% intervals. (b) Maximum arm joint velocity utilization, median and 90th percentile, against the actuator limit. (c) Peak hand linear speed, median and 90th percentile. The dashed line marks the maximum of the demonstrated speed range, $r = 2$, chosen without observing any learner outcome.

temporally ensembles predictions for the current action with normalized weights proportional to $\exp(-0.01i)$, where $i = 0$ indexes the oldest valid prediction. Table 7 collects the training and evaluation configuration.

B.3 ACT Checkpoint Selection

For UPRIGHT and PEG, ACT uses fixed episode-level train/validation partitions and normalization computed on training episodes only. Every 100 epochs, deterministic zero-latent validation ranks checkpoints by L1 averaged over valid action elements. We retain the five lowest-loss candidates plus the final checkpoint and evaluate each at nominal speed on a 100-episode development bank, with 10 mm planar jitter and no observation corruption or action noise. The selection order is greatest success count, lexicographically greatest task-stage vector, lowest deterministic validation L1, then earliest epoch. Checkpoint selection and final evaluation use disjoint episode banks. Expert and learner share indexed initial conditions within each final evaluation factor. Table 7 collects the partitions, seeds, and selected epochs, while Tables 9 and 10 report each replication.

B.4 Diffusion Policy and DAgger Training

To train Diffusion Policy (DP), we used all admitted demonstrations, a conditional one-dimensional U-Net with channel widths 256/512/1024, horizon 16, two observation steps, eight executed actions, and 100 diffusion steps for training and inference. Optimization uses AdamW, learning rate 10^{-4}, batch size 256, EMA power 0.75, and 300 epochs. The final epoch is fixed. DAgger uses an MLP with widths 512/256/128 and ELU activations, action MSE, Adam at 10^{-3}, and batch size 4096. An initial 40-epoch fit is followed by four rounds of 100 complete learner episodes and 40-epoch refits on the accumulated expert labels. Collection samples speedup factors uniformly over the task's demonstrated interval and adds Gaussian action noise of standard deviation 0.1 in half the environments. Final evaluation uses the round-four checkpoint with action noise disabled.

C Extended Performance Results

C.1 Complete Evaluation Grids

Complete success counts and UPRIGHT/PEG ACT replications are reported in Tables 8 to 10. Additional PARCEL ACT seeds are compared in Appendix C.2.

C.2 Parcel Speed Response and Training Seeds

Table 11 and Fig. 5 compare PARCEL ACT training seeds and demonstration subsets; Fig. 6 gives paired success-rate differences and episode outcomes. From $r = 1$ to $r = 2$, ACT-B and ACT-C success decreases by 0.34 and 0.48, compared with 0.16 for the expert, from nominal rates of 0.70,

Table 7: Training and evaluation configuration. ACT seeds change parameter initialization, episode permutations, and sampled chunk starts. The index j, starting at zero in ascending factor order, determines the random seed for sampling UPRIGHT/PEG initial conditions.

Setting	Value
UPRIGHT ACT train/validation episodes	252/63 (315 total)
PEG ACT train/validation episodes	260/65 (325 total)
UPRIGHT/PEG permutation / chunk-start seeds	20260904 / 20260905
UPRIGHT/PEG nominal development bank	Base seed 42001, 32 environments
UPRIGHT/PEG final evaluation bank	Initial-state sampling seed $73001 + 1000j$
PARCEL final evaluation bank	Base seed 12345
ACT training seeds (main figures / replications)	0 / 1, 2
PARCEL ACT labels / training seeds	A/B/C / 0/1/2
Diffusion Policy / DAgger model seeds	42 / 1
UPRIGHT selected epochs (seeds 0/1/2)	2000/1600/1300
PEG selected epochs (seeds 0/1/2)	1400/2000/1600
PARCEL ACT final training loss (A/B/C, epoch 2000)	0.0388 / 0.0403 / 0.0405
Paired bootstrap random generator	NumPy default generator, seed 0

Table 8: Parcel Success Counts Across the Final Grid, 100 Episodes per Cell. Mint marks nominal expert and ACT-A success; salmon marks ACT-A at $r = 2$.

Actor / r	0.5	1	1.5	2	2.25	2.5	3
Expert	100	100	99	84	80	29	0
ACT-A	84	100	91	53	36	0	0
Diffusion Policy	59	68	73	54	9	2	0
DAgger	2	3	1	3	1	0	0

0.62, and 1.00, respectively. Beyond the demonstrated range, the expert–ACT-A difference is 44 percentage points at $r = 2.25$ and 29 at $r = 2.5$; both fail all episodes at $r = 3$. Appendix D.4 examines the expert's decline, and Appendix C.3 gives the demonstration-subset construction.

C.3 DEMONSTRATION SCALING

For each training-set size, we retrain ACT-A with the same architecture, optimizer, batch construction, loss, temporal ensembling, 2000-epoch budget, and pseudorandom parameter initialization. To construct a subset of size n, we sort the 297 demonstrations by speedup factor and select evenly spaced ranks, so every subset spans the demonstrated speed range with a matching median value of r. Evaluation reuses the main evaluation's initial-condition draws at each speed. Table 12 reports the values corresponding to the gray curves in Fig. 5. Nominal success increases from 0.27 to 0.44 to 1.00 as the training set grows from 50 to 100 to 297 demonstrations. At $r = 2$, the corresponding values are 0.46, 0.44, and 0.53, with overlapping Wilson intervals. Success at $r = 2$ therefore shows no consistent increase across these three fits.

Table 9: Upright Success Counts Across the Final Grid, 200 Episodes per Cell. Mint marks the nominal expert–ACT comparison in Fig. 1; salmon marks ACT at $r = 2$.

Actor / r	0.5	0.75	1	1.25	1.5	1.75	2	2.5
Expert	176	191	185	194	193	184	105	2
ACT seed 0	155	182	194	199	198	194	24	0
ACT seed 1	105	179	194	200	198	199	98	0
ACT seed 2	157	153	182	173	177	161	68	0
Diffusion Policy	155	150	181	174	180	172	57	2
DAgger	1	0	1	0	1	0	2	4

Table 10: Peg Success Counts Across the Final Grid, 200 Episodes per Cell. Mint marks the nominal expert–ACT comparison in Fig. 1; salmon marks ACT at $r = 1.5$.

Actor / r	0.5	0.75	1	1.25	1.5	1.75	2	2.5
Expert	195	198	182	176	150	163	172	174
ACT seed 0	166	171	191	172	3	0	0	0
ACT seed 1	127	145	179	148	0	0	0	0
ACT seed 2	199	195	197	186	0	0	0	0
Diffusion Policy	93	99	95	120	25	0	0	0
DAgger	0	0	0	1	0	0	0	0

Table 11: **PARCEL Success Rates Within the Demonstrated Range.** Each cell uses 100 episodes. Decrease denotes the change from $r = 1$ to $r = 2$, omitted for floor-saturated DAgger. Diffusion Policy rates have no verified paired or matched-distribution expert comparison. Mint marks nominal equality between the expert and ACT-A; salmon marks ACT-A at $r = 2$. Other learners start below expert nominal success.

Policy	$r = 1$	$r = 1.5$	$r = 2$	Decrease
Scripted expert	1.00	0.99	0.84	0.16
ACT-A	1.00	0.91	0.53	0.47
ACT-B	0.70	0.58	0.36	0.34
ACT-C	0.62	0.44	0.14	0.48
Diffusion Policy	0.68	0.73	0.54	0.14
DAgger	0.03	0.01	0.03	—

C.4 BASELINES OVER THE FULL SPEED GRID

Fig. 7 plots the expert, Diffusion Policy, and DAgger across the full evaluation grid. Neither learner reaches the expert's recorded success count at $r = 1$. PARCEL Diffusion Policy has 68/100 initial-pose mismatches with the expert at each speedup factor. DAgger has 3/100 mismatches at $r = 2$ and 77/100 at $r = 2.5$. These affected conditions support standalone rates only, because full distribution matching is also unverified. DAgger completes at most 3 of 100 episodes at any evaluated speed, so sensitivity to execution speed cannot be distinguished at this success level.

D PARCEL FAILURE DIAGNOSTICS

D.1 STAGE OUTCOMES

In addition to the events in Table 4, a drop requires summed distal contact force at or below 0.5 N for 0.2 s after acquisition and before release, while the insertion predicate is false. A jam requires parcel–receptacle contact above 2 N for 0.2 s during insertion without satisfying the insertion predicate. Excessive relative motion denotes hand–parcel translation above 8 mm or rotation above 10 degrees from the acquisition pose, while held before release.

Terminal classification applies the following precedence after excluding task success. Failure to acquire comes first, followed by a recorded drop: during lift or earlier, during reorientation, during transfer or pre-insertion, or during insertion. Insertion-phase drops are jams if a jam was recorded and

Table 12: Task success of ACT-A policies trained with different numbers of demonstrations, 100 episodes per cell with Wilson 95% intervals, under identical evaluation draws across policies. Mint highlights the full 297-demonstration policy used for the main ACT-A results.

Condition	$r = 1$	$r = 1.5$	$r = 2$
ACT-A (n=50)	0.27 [0.19, 0.36]	0.23 [0.16, 0.32]	0.46 [0.37, 0.56]
ACT-A (n=100)	0.44 [0.35, 0.54]	0.65 [0.55, 0.74]	0.44 [0.35, 0.54]
ACT-A (n=297)	1.00 [0.96, 1.00]	0.91 [0.84, 0.95]	0.53 [0.43, 0.62]

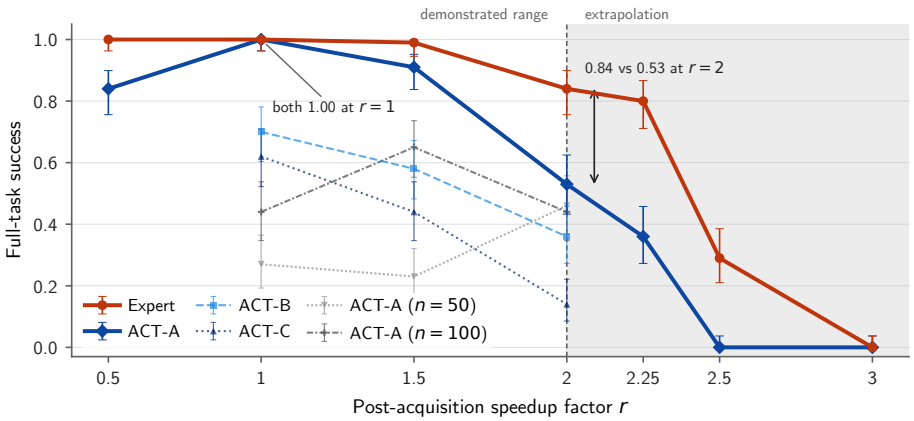
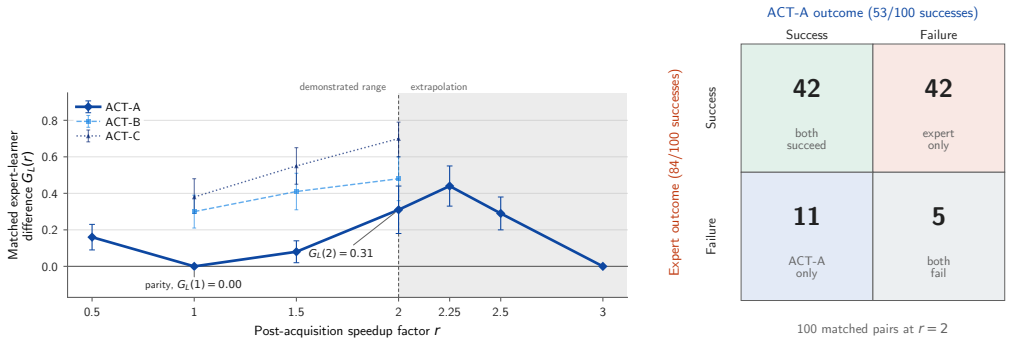


Figure 5: **PARCEL Success Across Speedup Factors and Training Seeds.** Pointwise Wilson 95% intervals use 100 episodes per condition. ACT-A/B/C denote three training runs; gray curves use ACT-A's configuration with 50 and 100 demonstrations. All three ACT seeds show larger success decreases from $r = 1$ to $r = 2$ than the expert. Shading marks extrapolation beyond demonstrated factors, $r > 2$.



(a) Matched success rate difference

(b) Matched outcomes at $r = 2$

Figure 6: **Paired Expert–ACT Comparisons.** (a) Expert-minus-learner success-rate difference $G_L(r)$, with paired bootstrap 95% intervals under the pairing assumptions in Sec. 4.2. ACT-A covers the full grid; ACT-B/C cover $r \in \{1, 1.5, 2\}$. (b) At $r = 2$, the expert and ACT-A both succeed on 42 episodes, the expert alone succeeds on 42, ACT-A alone succeeds on 11, and both fail on 5. Their marginal success counts are 84 and 53.

misalignments otherwise. If insertion never occurred, excessive relative motion before first receptacle contact takes precedence when the pre-insertion event occurred; failure to reach pre-insertion is a transfer failure; a recorded jam or excessive relative motion after receptacle contact is an insertion jam; all other cases are insertion misalignments. After an insertion event, failure to release is a release failure, failure to settle is a timeout, final loss of containment is other, and final orientation error above 10 degrees is an insertion misalignment. Earlier applicable categories take precedence over later ones.

D.1.1 FAILURE LOCALIZATION

Fig. 8a shows the stage counts underlying Sec. 5.4, and Fig. 9a shows terminal outcomes. At $r = 2$, the expert–ACT-A difference grows from 6 percentage points at pre-insertion to 31 at terminal success; insertion misalignments and jams account for 45 of ACT-A's 47 failures.

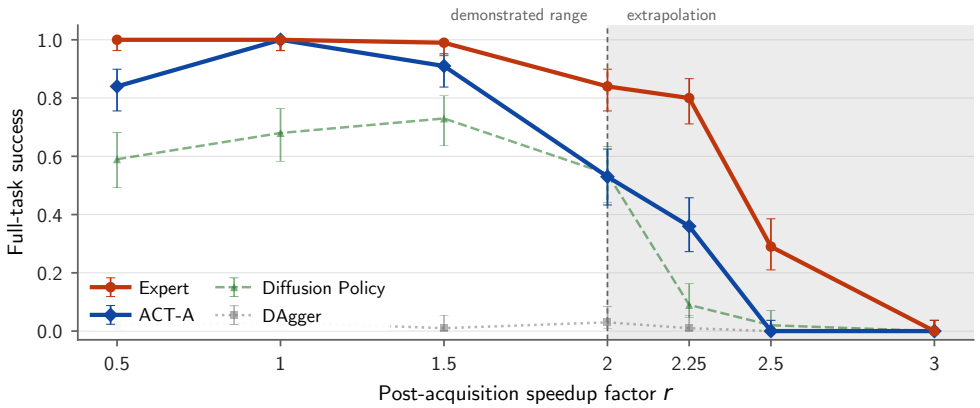


Figure 7: **PARCEL Baseline Task Success Across the Full Speed Grid.** Task success against the speedup factor for the phases after acquisition for the expert, ACT-A, Diffusion Policy, and DAgger, with pointwise Wilson 95% intervals for each policy over 100 episodes per point. These are standalone curves, not paired baseline differences. The shaded band marks extrapolation beyond the demonstrated range, $r > 2$.

D.2 RELATIVE-MOTION HANDOFF

The relative-motion handoff preserves the physical state when each actor's downstream commands are replaced. For an episode of actor π at speedup factor r, let t_H be the first control step after acquisition at which the parcel center has risen at least 40 mm from its initial height, provided the release phase has not begun. Episodes without this trigger retain their original controller. Write $\phi_r(t)$ for the speed-dependent phase clock induced by Eq. (5); this clock continues through the handoff. Let $T_H^{W,E}(\phi) \in \mathrm{SE}(3)$ denote the forward-kinematic hand pose of the expert's nominal joint command at phase ϕ, in world frame W. For this episode, the relative hand-pose command is

$$\Delta T^E_{H,\pi}(t) = \left[T_H^{W,E}(\phi_r(t_H))\right]^{-1} T_H^{W,E}(\phi_r(t))\,, \qquad t \geq t_H. \tag{9}$$

Let $A^W_\pi \in \mathrm{SE}(3)$ be the forward-kinematic hand pose of the evaluated policy's arm command at handoff after subtracting its static servo offset. The shared controller computes the target

$$T^{W,d}_{H,\pi}(t) = A^W_\pi \, \Delta T^E_{H,\pi}(t), \tag{10}$$

where the superscript d marks the commanded target. The static servo offset is the commanded minus measured arm-joint configuration at the end of the grasp dwell, clamped to the controller's correction limits. A shared damped least-squares controller tracks this target and initializes its integral correction with that offset to maintain command continuity. The hand-joint targets at handoff are held fixed until the scheduled release, when the controller applies the expert's opening commands.

Parcel retention is evaluated from handoff until insertion begins or the parcel first contacts the receptacle; relative motion uses the hand–parcel transform at handoff as its reference. Task completion is evaluated by continuing the controller through insertion, release, and settling. Because the handoff preserves actor-specific physical states, completion differences do not isolate state error from subsequent contact effects.

D.2.1 RETENTION AND TASK COMPLETION

Across $r \in \{1, 1.5, 2\}$, the expert, ACT-A, and ACT-C each acquire and retain all 300 parcels to the free-space endpoint. ACT-B acquires 100, 98, and 100 parcels and retains 98, 94, and 99, respectively (Fig. 8b). Median relative translation stays below 5.8 mm and rotation below $5.4°$ across actors and factors. Continuing through insertion, release, and settling at $r = 2$ yields 95/100 successes after expert acquisition and 64/100, 55/100, and 15/100 after ACT-A/B/C acquisition. ACT-A thus completes 64 handoff episodes versus 53 under its own downstream commands. ACT-C contacts the receptacle before the nominal pre-insertion endpoint in 65% of episodes. Failed ACT-A handoffs have larger median pose offsets, with substantial overlap between outcome groups (Fig. 9b).

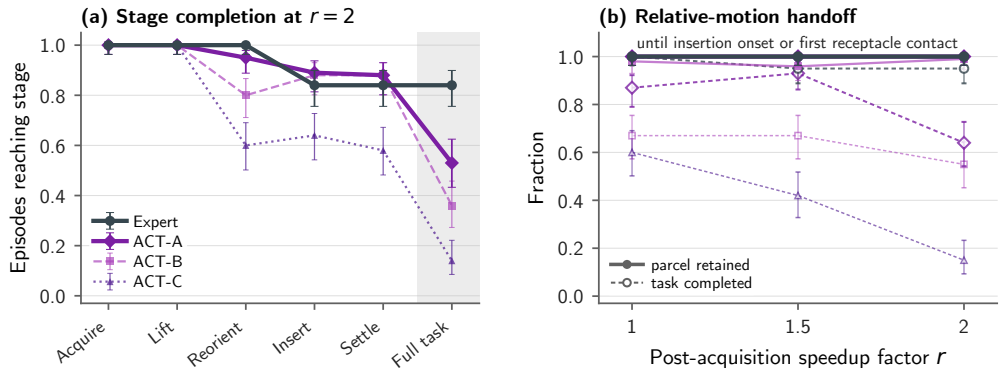
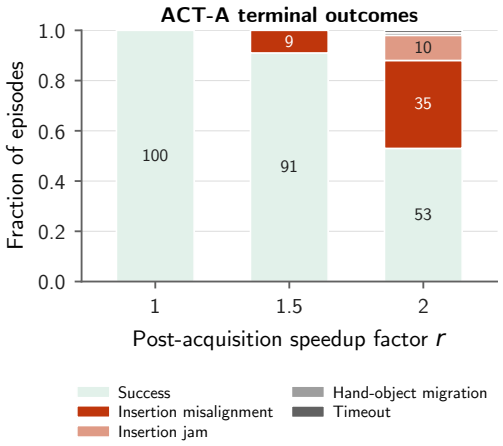
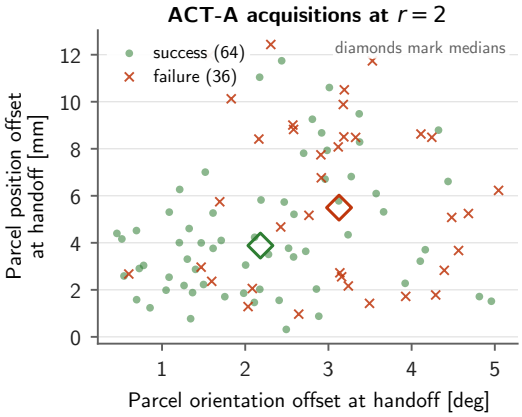


Figure 8: **Stage Completion and Relative-Motion Handoff.** (a) Fraction of episodes in which each event occurs at $r = 2$ for the expert and the three ACT policies. The expert and ACT-A coincide at acquisition and lift, while their terminal success rates differ. Each value is a marginal event frequency. (b) Solid curves report the fraction of triggered handoffs retaining the parcel until insertion onset or first receptacle contact under Eq. (10). Dashed curves report task success over all 100 episodes per actor and factor, including episodes without handoff. For ACT-B at $r = 1.5$, the denominators differ: retention is 94/98 handoffs and task success is 67/100 episodes.



(a) **Terminal Outcomes Across Demonstrated Speeds.** Outcome composition for ACT-A at $r \in \{1, 1.5, 2\}$.



(b) **Handoff Pose Offset and Downstream Outcome at $r = 2$.** ACT-A parcel-pose offsets relative to the matched expert at handoff.

Figure 9: **ACT-A Failure Localization and Diagnostics of the State at Handoff.** (a) Normalized terminal outcomes show 100 successes at $r = 1$, nine insertion misalignments at $r = 1.5$, and at $r = 2$ a composition of 53 successes, 35 insertion misalignments, 10 insertion jams, one failure classified as excessive parcel slip, and one timeout. The increasing failure rate is therefore dominated by insertion failures. (b) At $r = 2$, failed handoffs have larger median rotational and translational offsets: 3.1° and 5.5 mm for failures, compared with 2.2° and 3.9 mm for successes. The groups overlap substantially.

D.3 Hand–Parcel Relative Motion

Denoting $T^W_{H,t}$ and $T^W_{O,t}$ as the hand and object poses in the world frame, the object pose in the hand frame is $T^H_{O,t} = (T^W_{H,t})^{-1} T^W_{O,t}$. We let t_G denote the first step satisfying the acquisition predicate in Table 4 and write the hand–parcel relative motion as

$$\Xi_{HO}(t) = \left(T^H_{O,t_G}\right)^{-1} T^H_{O,t}, \tag{11}$$

where $R_\Xi(t)$ is the rotation block of $\Xi_{HO}(t)$. Let $\mathcal{T}_G$ contain the recorded control steps at or after acquisition, before scheduled release, while the parcel has not previously been marked as dropped.

We report the peak hand–parcel relative translation (d_p) and rotation (d_R), given respectively as the maximum Euclidean norm of the translation vector and the maximum angle of rotation in axis-angle representation over $\mathcal{T}_G$:

$$d_p = \max_{t \in \mathcal{T}_G} \|\mathrm{trans}(\Xi_{HO}(t))\|_2 \,, \qquad d_R = \max_{t \in \mathcal{T}_G} \left\|\log(R_\Xi(t))^\vee\right\|_2 \,. \tag{12}$$

Here, trans extracts the translation vector, and $(\log R)^\vee$ is the axis-angle vector of rotation R. Fig. 10 reports these quantities for the expert and ACT-A across speedup factors.

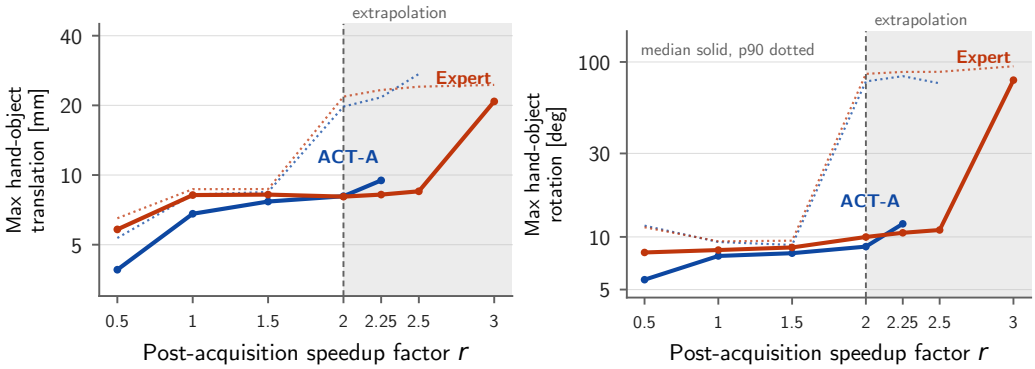


Figure 10: **Hand-Parcel Relative Motion Across Execution Speeds.** Median and 90th-percentile peak hand–parcel relative translation (left) and rotation (right) for the expert and ACT-A. Translation and rotation are limited across the demonstrated range $r \leq 2$, despite the substantial expert–learner difference in task success at $r = 2$. The shaded region denotes execution speeds outside the demonstrated range.

D.4 Joint Velocity Utilization and Expert Diagnostics

At $r = 2$, arm joint velocity utilization (Eq. (8)) has medians of 0.108 for the expert and 0.107 for ACT-A, with 90th percentiles of 0.111 and 0.117. These values are well below the velocity limit. Between $r = 2.25$ and $r = 2.5$, the expert's median minimum orientation error before insertion crosses the 10° tolerance as success falls from 0.80 to 0.29. Median arm joint velocity utilization stays below 0.14 through $r = 2.5$. Table 13 gives the orientation, insertion-depth, contact-force, and motion statistics.

Table 13: Expert success proportions and median diagnostic quantities over 100 episodes per speedup factor. Mint marks uniform success and salmon complete failure. Bold indicates speeds at which the minimum orientation error before insertion exceeds the 10° tolerance.

r	0.5	1	1.5	2	2.25	2.5	3
Success	1.00	1.00	0.99	0.84	0.80	0.29	0.00
Min. Pre-Insertion Orientation Error [°]	2.8	2.2	4.7	7.6	8.9	**10.3**	**10.8**
Insertion Depth [mm]	63.7	56.3	52.5	50.0	49.5	49.1	30.4
Peak Receptacle Contact Force [N]	2.7	3.1	4.7	5.7	7.0	8.1	13.6
Arm Joint Velocity Utilization	0.10	0.12	0.08	0.11	0.12	0.13	0.16
Peak Hand–Parcel Translation [mm]	5.8	8.2	8.2	8.1	8.2	8.5	20.8

D.5 Realized Contact Margins

Force closure means that feasible contact forces can balance an arbitrary external wrench. At acquisition, at the end of reorientation, and at insertion start, we score the realized fingertip contact set with the Ferrari–Canny margin ε (Ferrari & Canny, 1992). Contacts with measured force above 1 N supply a point p and an inward normal. At nominal friction $\mu = 0.5$, each contact contributes eight unit-norm friction-cone edge forces f and their wrenches $w = (f, (p - c) \times f)$, where c is the parcel center. Positions are expressed in metres, with no additional torque scaling. For a wrench convex

hull containing the origin, ε is the minimum distance from the origin to a hull facet; non-containing or degenerate hulls receive the recorded value -1. Within this contact model, $\varepsilon > 0$ denotes force closure and $\varepsilon \leq 0$ its absence.

The margin describes the contacts realized in simulation, not a planned or idealized grasp. The condition $\varepsilon > 0$ has a physically defined threshold and zero fitted parameters, and we evaluate it as a candidate necessary condition for downstream success.

D.5.1 FORCE CLOSURE AND DOWNSTREAM SUCCESS

Across the six PARCEL actors (expert, ACT-A/B/C, Diffusion Policy, and DAgger) and their evaluated speedup factors, none of the 414 acquired episodes with $\varepsilon \leq 0$ at acquisition completes the task. All 700 expert acquisitions satisfy force closure. The same empirical implication holds for the two checkpoints trained with different numbers of demonstrations: the additional 79 acquisitions without force closure all fail. The corresponding counts among acquired episodes are 42/542 for ACT-A (7.75%), 63/298 for ACT-B (21.1%), and 102/300 for ACT-C (34.0%).